

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

## **IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

PAT-NO: JP408297626A  
DOCUMENT- JP 08297626 A  
IDENTIFIER:  
TITLE: NETWORK INTERFACE AND METHOD FOR PROCESSING PACKET  
IN NETWORK INTERFACE  
PUBN-DATE: November 12, 1996

INVENTOR-INFORMATION:

| NAME            | COUNTRY |
|-----------------|---------|
| RANDY, B OSBORN |         |

ASSIGNEE-INFORMATION:

| NAME                                | COUNTRY |
|-------------------------------------|---------|
| MITSUBISHI ELECTRIC RES LAB INC N/A |         |

APPL-NO: JP08022600  
APPL-DATE: February 8, 1996

INT-CL (IPC): G06F013/12 , G06F013/00

ABSTRACT:

PROBLEM TO BE SOLVED: To allow an application to control where data should be stored to by performing access to one part of an outside list according to an outside empty buffer list of each channel, and the predicated size of an inputted packet.

SOLUTION: When a packet is received, the packet is selected from a list, and data are not copied to the buffer of an operating system level but stored in the buffer. That is, each application 46 and 48 is provided with independent empty buffer lists 60 and 62, and they can be stored in an application level memory 42, operating system 38, or network interface 31. The network interface 31 specifies the next buffer in which data from the received packet should be stored for each packet. Thus, the application can control the empty buffer in which the data should be stored.

(11)特許出願公開番号

(43)公開日 平成8年(1996)11月12日

| (51) Int.Cl. <sup>8</sup> | 識別記号  | 片内整理番号  | F I           | 技術表示箇所  |
|---------------------------|-------|---------|---------------|---------|
| G 0 6 F 13/12             | 3 3 0 | 7368-5E | G 0 6 F 13/12 | 3 3 0 G |
| 13/00                     | 3 5 3 | 7368-5E | 13/00         | 3 5 3 Q |

審査請求 有 請求項の数20 O L (全 11 頁)

|             |                |         |   |
|-------------|----------------|---------|---|
| (21)出願番号    | 特願平8-22600     | (71)出願人 | 595151497<br>ミツビシ・エレクトリック・リサーチ・ラボラトリーズ・インコーポレイテッド<br>MITSUBISHI ELECTRIC<br>RESEARCH LABORATORIES, INC.<br>アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ブロードウェイ 201 |
| (22)出願日     | 平成8年(1996)2月8日 | (72)発明者 | ランディ・ビー・オズボルン<br>アメリカ合衆国、マサチューセッツ州、ケンブリッジ、パーカー・ストリート 39   |
| (31)優先権主張番号 | 08/422344      | (74)代理人 | 弁理士 曾我 道照 (外6名)   |
| (32)優先日     | 1995年4月14日     |         |   |
| (33)優先権主張国  | 米国(US)         |         |   |

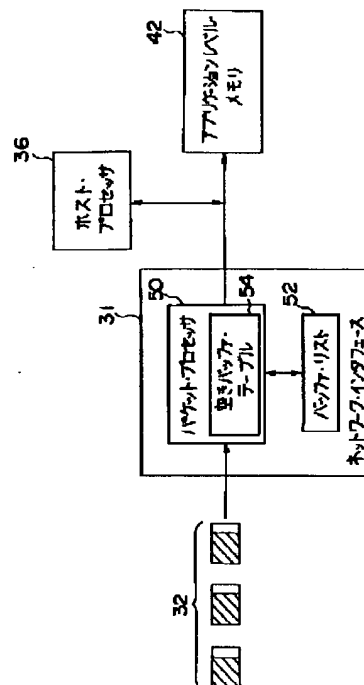
最終頁に続く

(54) 【発明の名称】 ネットワーク・インタフェースおよびネットワーク・インタフェースにおけるパケット処理方法

(57) 【要約】

【課題】 広域的バッファ・キューは多重通信チャネル相互間で共用されるのでパケットが割当てられるバッファを事前に知ることはできず、データが実際に転送される記憶位置をアプリケーションが制御できず、従ってある種のアプリケーションデータ構造で占められたメモリ・ページにパケット・データを直接記憶することができなかった。

【解決手段】 チャネル毎の外部空きバッファ・リストと、入力されるパケットの予測サイズに応じて外部リストの一部分をアクセスすることによって、各々の入力されるパケット用にロードされた、外部の空きバッファ・リストに対応するチャネル毎の内部の空きバッファ・リスト54とを利用するネットワーク・インタフェース31を構成し、アプリケーションの入力パケットの利用バッファの制御を可能にする。



## 【特許請求の範囲】

【請求項1】 データとチャネルを示す見出し情報とを含むパケットを処理するネットワーク・インタフェースにおいて、

各チャネル毎に、該チャネル用のメモリ内の利用できる空きバッファ表示を記憶するためのバッファ・リスト・メモリと、および、

パケットを受信する入力を受信するパケット・プロセッサであって、

パケットを受信すると動作して、パケットの見出し情報によって表示されるチャネルについて、パケットの予測サイズに従って数が選択される多数の空きバッファ表示を検索するためにバッファ・リスト・メモリをアクセスし、かつ検索した表示内容を内部メモリに記憶するための手段、

パケットを受信すると動作して、検索されたバッファ・リストから使用するバッファを特定するための手段、パケット内のデータを特定されたバッファへと転送する手段、

とを含むパケット・プロセッサと、から構成したことを特徴とするネットワーク・インタフェース。

【請求項2】 各チャネル毎に上記バッファ・リスト・メモリに記憶された空きバッファの表示が、複数の空きバッファの表示リストのリストとして編成され、各リストの長さが入力されるパケットの予測サイズに対応することを特徴とする請求項1に記載のネットワーク・インタフェース。

【請求項3】 空きバッファの表示の各リストが次のバッファ・リストの上記バッファ・リスト・メモリ内の記憶位置の表示を含むことを特徴とする請求項2に記載のネットワーク・インタフェース。

【請求項4】 各チャネル毎の空きバッファ表示と関連する、上記バッファ・リスト・メモリ内の割り込みコードをさらに備え、

上記パケット・プロセッサが、割り込みコードにตอบสนองして動作し、さらにバッファ・リストが必要であることを表示する割り込みコードを作成する手段を含んだことを特徴とする請求項1に記載のネットワーク・インタフェース。

【請求項5】 上記パケット・プロセッサをアクセスできるプロセッサが割り込みにตอบสนองして空きバッファに関する表示をさらに作成することを特徴とする請求項4に記載のネットワーク・インタフェース。

【請求項6】 上記プロセッサで実行されるアプリケーションプログラムが割り込みにตอบสนองして空きリスト表示をさらに作成することを特徴とする請求項5に記載のネットワーク・インタフェース。

【請求項7】 空きバッファを表示する各リストが割り込みコードを含むことを特徴とする請求項4に記載のネットワーク・インタフェース。

【請求項8】 オペレーティング・システムが割り込みにตอบสนองして、空きバッファ表示をさらに作成することを特徴とする請求項4に記載のネットワーク・インタフェース。

【請求項9】 アプリケーションプログラムが割り込みにตอบสนองして空きバッファ表示をさらに作成することを特徴とする請求項4に記載のネットワーク・インタフェース。

【請求項10】 ネットワーク・インタフェースを非同期転送モードのネットワークに接続したことを特徴とする請求項1に記載のネットワーク・インタフェース。

【請求項11】 特定されたバッファが使用され、データがパケット内に残っていると動作して、検索されたリスト中の別のバッファを特定し、パケットから別のバッファへのデータの転送を継続するための手段をさらに上記パケット・プロセッサ内に備えたことを特徴とする請求項1に記載のネットワーク・インタフェース。

【請求項12】 パケットが検索されると動作して、空きバッファ表示が外部バッファ・リスト・メモリ内に記憶されていることを確認するための手段をさらに上記パケット・プロセッサ内に備えたことを特徴とする請求項1に記載のネットワーク・インタフェース。

【請求項13】 上記確認手段が、アプリケーションへの割り込みコードを発生し、空きバッファ・リストを要求する手段を含むことを特徴とする請求項12に記載のネットワーク・インタフェース。

【請求項14】 空きバッファ表示の一部分を検索する手段が、空きバッファ表示の次の一部分の表示内容を検索する手段を含んだことを特徴とする請求項1に記載のネットワーク・インタフェース。

【請求項15】 空きバッファ表示の次の一部分の表示がない場合は、次の一部分の表示を検索する上記手段が、パケット・プロセッサにアクセスできるプロセッサに、空きバッファのより多くの表示が作成され、バッファ・リスト・メモリに記憶されるように要求する割り込みコードを作成することを特徴とする請求項14に記載のネットワーク・インタフェース。

【請求項16】 上記パケット・プロセッサにアクセスできるプロセッサが、アプリケーションが必要とするデータ蓄積位置に応じて空きバッファ表示リストに特定のバッファ記憶位置を事前装填することによって、プロセッサがデータ蓄積位置を制御することを特徴とする請求項1に記載のネットワーク・インタフェース。

【請求項17】 データとチャネルを示す見出し情報とを含むパケットを処理するネットワーク・インタフェースにおいて、

チャネル毎に受信されるパケットからの情報を記憶する空きバッファ表示を記憶するための、各チャネル毎の内部バッファ・リスト・メモリを含むパケット・プロセッサと、

各チャネルごとに使用される空きバッファ表示のリストを含む外部バッファ・リスト・メモリと、  
上記外部バッファ・リスト・メモリ内の空きバッファ表示のリストの一部分をアクセスし、かつパケット・プロセッサの内部バッファ・リストへと表示を複写するための、上記パケット・プロセッサ内の手段と、  
を備え、上記部分が入力されるパケットの予測サイズに応じてアクセスされることを特徴とするネットワーク・インタフェース。

【請求項18】 データとチャネルを示す見出し情報とを含むパケットを処理するネットワーク・インタフェースであって、  
各チャネル毎に、該チャネル用のメモリ内の利用できる空きバッファ表示を記憶するためのバッファ・リスト・メモリであり、このバッファ・リスト・メモリ内に記憶された各チャネル毎の空きバッファの表示が複数の空きバッファ表示リストのリストとして編成され、各リストの長さが入力されるパケットの予測サイズに対応しているものと、およびパケットを受信する入力を受けるパケット・プロセッサであり、  
パケットを受信すると動作して、パケットが表示するチャネル用に空きバッファの表示リストの一つを検索するために上記バッファ・リスト・メモリをアクセスし、かつ検索した表示を内部メモリに記憶する手段、  
パケットを受信すると動作して、検索されたバッファ・リストから使用するバッファを特定するための手段、  
パケット内のデータを特定されたバッファへと転送する手段、を含むものと、  
から構成したことを特徴とするネットワーク・インタフェース。

【請求項19】 データとチャネルを示す見出し情報とを含むパケットを処理するネットワーク・インタフェースであって、  
各チャネル毎に、該チャネル用のメモリ内の利用できる空きバッファ表示と、各チャネル毎の空きバッファ表示と関連する、バッファ・リスト・メモリ内の割り込みコードとを記憶するためのバッファ・リスト・メモリと、  
およびパケットを受信する入力を受けるパケット・プロセッサであり、  
パケットを受信すると動作して、パケットが表示するチャネル用に空きバッファの、パケットの予測サイズに従って数が選択される多数の表示リストを検索するために上記バッファ・リスト・メモリをアクセスし、かつ検索した表示を内部メモリに記憶する手段、  
パケットを受信すると動作して、より多くのバッファ・リストが必要であることを表示する割り込みを発生する手段、  
パケットを受信すると動作して、検索したバッファ・リストから使用されるバッファを特定する手段、  
パケット内のデータを特定されたバッファへと転送する

手段、  
とを含むパケット・プロセッサと、  
から構成したことを特徴とするネットワーク・インタフェース。

【請求項20】 データとチャネルを示す見出し情報とを含むパケットをネットワーク・インタフェースによって処理する方法であって、  
第1メモリに各通信チャネル毎の複数の空きバッファ表示を記憶するステップと、  
10 パケットを受信するステップと、  
見出し情報からパケットの通信チャネルを決定するステップと、  
パケットの予測サイズに応じて決定されたチャネル用にメモリから空きバッファ情報の一部分を検索し、その部分を第2メモリに記憶するステップと、  
第2メモリ内の検索された表示から使用されるバッファを選択するステップと、  
パケットからのデータを第2メモリ内に記憶するステップと、  
20 からなることを特徴とするネットワーク・インタフェースにおけるパケット処理方法。

#### 【発明の詳細な説明】

##### 【0001】

【発明の属する技術分野】この発明はコンピュータ・システム用のネットワーク・インタフェース、特に、ネットワーク・インタフェースが別のコンピュータからパケットで受信したデータを蓄積する空きバッファ・リストの管理に関する。

##### 【0002】

30 【従来の技術】ネットワーク化されたコンピュータ・システムでは、コンピュータは一般的には別のコンピュータと通信するためにネットワーク・インタフェースを介してネットワークに接続される。データ・パケットを受信するために、コンピュータには空いたバッファ領域、すなわちそこにデータを記憶できる空いた記憶位置が必要である。一般にコンピュータは、設立された全ての通信チャネル間で共用される、これまで広域的空きバッファ待ち行列(キュー)もしくはリストと呼ばれてきた単一の空きバッファ待ち行列(キュー)を有している。

40 【0003】異なる通信チャネルを使用するアプリケーション間の非干渉性、すなわち保護を確実にするために、一人のユーザーの複数のアプリケーション、または複数のユーザー相互間のアプリケーションのいずれの場合でも、広域的空きバッファ待ち行列(キュー)は信託された構成要素(trusted entity)によって管理される。通常は、この信託された構成要素はオペレーティング・システムであり、広域的空きバッファ待ち行列(キュー)はオペレーティング・システム専用のメモリ・スペースから割当てられる。

50 【0004】広域的空きバッファ待ち行列(キュー)を管

理するオペレーティング・システムの動作は次のとおりである。最初に、パケットが到着すると、広域的空きバッファ・リストからバッファが取り出され、それにパケット・データが充填される。パケット・データに追加のバッファが必要な場合は、次にバッファが広域的空きバッファ待ち行列(キュー)から取り出され、データが充填された後、そのパケット用に以前に充填されたバッファへと連結される。アプリケーションへのポーリング、ブロッキングまたは割り込みを介してそのデータが利用できるという表示がなされた後、アプリケーションはオペレーティング・システムの呼び出しを行い、それによってデータは2つの方式の一方を利用して適応業務スペースへと転送される。

【0005】第1の方式では、オペレーティング・システムはアプリケーションによって指定されたアプリケーションスペースの記憶位置にデータを複写する。第2の方式では、オペレーティング・システムはバッファを含むメモリ・ページをアプリケーションスペースのアプリケーションによって指定された記憶位置へとマッピングする。バッファは一般に、このようなマッピング効果を得るためにメモリ・ページのサイズと等しく、また、それと位置合わせされて構成されている。アプリケーションが、オペレーティング・システムが完全な受信パケットをアプリケーションに転送することを呼び出すように受信ハンドラを設定することも可能である。この場合、オペレーティング・システムはデータを転送するアドレスを搬送することも要求される。

【0006】これらの方式には双方ともオペレーティング・システム呼び出しの高いコストが伴う。勿論、このような呼び出しの経費は大量のデータ転送の場合は容易に償還できる。第1の方式ではデータを複写するコストが高いが、アプリケーションが望むアプリケーションスペースにデータを正確に配することができるという利点を有している。第2の方式は直接的にはメモリ・マッピング表を処理するコストと、おそらくは結果的に生ずるわずかなメモリ転移ミスが伴うだけである。しかし、バッファを含むメモリ・ページはバッファ以外のどのアプリケーションデータにも使用できず、それによって、アプリケーションが望む位置にデータを正確に配することが間接的に制約される。

【0007】

【発明が解決しようとする課題】以上のような従来のネットワーク・インタフェースでは、広域的バッファ待ち行列(キュー)は多重の通信チャネル相互間で共用されるので、どのパケットがどのバッファに割当てられるかを前もって知ることはできない。従って、データが実際に転送される記憶位置をアプリケーションが制御する方法はない。このように制御がなされないことは、ある種のアプリケーションデータ構造で占められたメモリ・ページにパケット・データを直接記憶することができないこ

とを意味している。その結果、データを複写する必要がある。この方式は、アプリケーションが大量のデータ・ブロックを処理するものと想定されている場合に、大量のデータを配達することを指向するものである。

【0008】オペレーティング・システム呼び出しのコストによって、双方の方式とも少量のメッセージ用の処理時間の経費が比較的高く、待ち時間のメッセージを送る要求の度合いが低いアプリケーションでは使用ににくい。さらに、別のアプリケーション構造と混用されたメモリ・ページのデータをアプリケーションが必要とすることがあるため、双方の方式に、すなわち第1の方式では直接的に、また、第2の方式では間接的に複写のコストが加わる。従来のネットワーク・インタフェースでは以上のような課題があった。

【0009】この発明は、チャネル毎の外部の空きバッファ・リストと、入力されるパケットの予測サイズに応じて外部リストの一部分をアクセスすることによって、各々の入力されるパケット用にロードされた対応するチャネル毎の内部の空きバッファ・リストと、を利用するネットワーク・インタフェースおよびネットワーク・インタフェースにおけるパケット処理方法を提供することによって、アプリケーションがどこにデータを蓄積するかを制御できる等、従来技術に伴う上記のような問題点を克服することを目的とする。

【0010】

【課題を解決するための手段】上記の目的に鑑み、この発明の第1の発明は、データとチャネルを示す見出し情報とを含むパケットを処理するネットワーク・インタフェースにおいて、各チャネル毎に、該チャネル用のメモリ内の利用できる空きバッファ表示を記憶するためのバッファ・リスト・メモリと、および、パケットを受信する入力を受信するパケット・プロセッサであって、パケットを受信すると動作して、パケットの見出し情報によって表示されるチャネルについて、パケットの予測サイズに従って数が選択される多数の空きバッファ表示を検索するために上記バッファ・リスト・メモリをアクセスし、かつ検索した表示内容を内部メモリに記憶するための手段、パケットを受信すると動作して、検索されたバッファ・リストから使用するバッファを特定するための手段、パケット内のデータを特定されたバッファへと転送する手段、とを含むパケット・プロセッサと、から構成したことを特徴とするネットワーク・インタフェースにある。

【0011】この発明の第2の発明は、各チャネル毎に上記バッファ・リスト・メモリに記憶された空きバッファの表示が、複数の空きバッファの表示リストのリストとして編成され、各リストの長さが入力されるパケットの予測サイズに対応することを特徴とする請求項1に記載のネットワーク・インタフェースにある。

【0012】この発明の第3の発明は、空きバッファの

7

表示の各リストが次のバッファ・リストの上記バッファ・リスト・メモリ内の記憶位置の表示を含むことを特徴とする請求項2に記載のネットワーク・インタフェースにある。

【0013】この発明の第4の発明は、各チャネル毎の空きバッファ表示と関連する、上記バッファ・リスト・メモリ内の割り込みコードをさらに備え、上記バケット・プロセッサが、割り込みコードにตอบสนองして動作し、さらにバッファ・リストが必要であることを表示する割り込みコードを作成する手段を含んだことを特徴とする請求項1に記載のネットワーク・インタフェースにある。

【0014】この発明の第5の発明は、上記バケット・プロセッサにアクセスできるプロセッサが割り込みにตอบสนองして空きバッファに関する表示をさらに作成することを特徴とする請求項4に記載のネットワーク・インタフェースにある。

【0015】この発明の第6の発明は、上記プロセッサで実行されるアプリケーションプログラムが割り込みにตอบสนองして空きリスト表示をさらに作成することを特徴とする請求項5に記載のネットワーク・インタフェースにある。

【0016】この発明の第7の発明は、空きバッファを表示する各リストが割り込みコードを含むことを特徴とする請求項4に記載のネットワーク・インタフェースにある。

【0017】この発明の第8の発明は、オペレーティング・システムが割り込みにตอบสนองして、空きバッファ表示をさらに作成することを特徴とする請求項4に記載のネットワーク・インタフェース。

【0018】この発明の第9の発明は、アプリケーションプログラムが割り込みにตอบสนองして空きバッファ表示をさらに作成することを特徴とする請求項4に記載のネットワーク・インタフェースにある。

【0019】この発明の第10の発明は、ネットワーク・インタフェースを非同期転送モードのネットワークに接続したことを特徴とする請求項1に記載のネットワーク・インタフェース。

【0020】この発明の第11の発明は、特定されたバッファが使用され、データがバケット内に残っていると動作して、検索されたリスト中の別のバッファを特定し、バケットから別のバッファへのデータの転送を継続するための手段をさらに上記バケット・プロセッサ内に備えたことを特徴とする請求項1に記載のネットワーク・インタフェースにある。

【0021】この発明の第12の発明は、バケットが検索されると動作して、空きバッファ表示が外部バッファ・リスト・メモリ内に記憶されていることを確認するための手段をさらに上記バケット・プロセッサ内に備えたことを特徴とする請求項1に記載のネットワーク・インタフェースにある。

8

【0022】この発明の第13の発明は、上記確認手段が、アプリケーションへの割り込みコードを発生し、空きバッファ・リストを要求する手段を含むことを特徴とする請求項12に記載のネットワーク・インタフェースにある。

【0023】この発明の第14の発明は、空きバッファ表示の一部分を検索する手段が、空きバッファ表示の次の一部分の表示内容を検索する手段を含んだことを特徴とする請求項1に記載のネットワーク・インタフェースにある。

【0024】この発明の第15の発明は、空きバッファ表示の次の一部分の表示がない場合は、次の一部分の表示を検索する上記手段が、バケット・プロセッサにアクセスできるプロセッサに、空きバッファのより多くの表示が作成され、バッファ・リスト・メモリに記憶されるように要求する割り込みコードを作成することを特徴とする請求項14に記載のネットワーク・インタフェースにある。

【0025】この発明の第16の発明は、上記バケット・プロセッサにアクセスできるプロセッサが、アプリケーションが必要とするデータ蓄積位置に応じて空きバッファ表示リストに特定のバッファ記憶位置を事前装填することによって、プロセッサがデータ蓄積位置を制御することを特徴とする請求項1に記載のネットワーク・インタフェースにある。

【0026】この発明の第17の発明は、データとチャネルを示す見出し情報とを含むバケットを処理するネットワーク・インタフェースにおいて、チャネル毎に受信されるバケットからの情報を記憶する空きバッファ表示を記憶するための、各チャネル毎の内部バッファ・リスト・メモリを含むバケット・プロセッサと、各チャネルごとに使用される空きバッファ表示のリストを含む外部バッファ・リスト・メモリと、上記外部バッファ・リスト・メモリ内の空きバッファ表示のリストの一部分にアクセスし、かつバケット・プロセッサの内部バッファ・リストへと表示を複写するための、上記バケット・プロセッサ内の手段と、を備え、上記部分が入力されるバケットの予測サイズに応じてアクセスされることを特徴とするネットワーク・インタフェースにある。

【0027】この発明の第18の発明は、データとチャネルを示す見出し情報とを含むバケットを処理するネットワーク・インタフェースであって、各チャネル毎に、該チャネル用のメモリ内の利用できる空きバッファ表示を記憶するためのバッファ・リスト・メモリであり、このバッファ・リスト・メモリ内に記憶された各チャネル毎の空きバッファの表示が複数の空きバッファ表示リストのリストとして編成され、各リストの長さが入力されるバケットの予測サイズに対応しているものと、および、バケットを受信する入力を受信するバケット・プロセッサであり、バケットを受信すると動作して、バケット

が表示するチャンネル用に空きバッファの表示リストの一つを検索するために上記バッファ・リスト・メモリをアクセスし、かつ検索した表示を内部メモリに記憶する手段、パケットを受信すると動作して、検索されたバッファ・リストから使用するバッファを特定するための手段、パケット内のデータを特定されたバッファへと転送する手段、を含むものと、から構成したことを特徴とするネットワーク・インタフェースにある。

【0028】この発明の第19の発明は、データとチャンネルを示す見出し情報とを含むパケットを処理するネットワーク・インタフェースであって、各チャンネル毎に、該チャンネル用のメモリ内の利用できる空きバッファ表示と、各チャンネル毎の空きバッファ表示と関連する、バッファ・リスト・メモリ内の割り込みコードとを記憶するためのバッファ・リスト・メモリと、および、パケットを受信する入力を受けるパケット・プロセッサであり、パケットを受信すると動作して、パケットが表示するチャンネル用に空きバッファの、パケットの予測サイズに従って数が選択される多数の表示リストを検索するために上記バッファ・リスト・メモリをアクセスし、かつ検索した表示を内部メモリに記憶する手段、パケットを受信すると動作して、より多くのバッファ・リストが必要であることを表示する割り込みを発生する手段、パケットを受信すると動作して、検索したバッファ・リストから使用されるバッファを特定する手段、パケット内のデータを特定されたバッファへと転送する手段、を含むパケット・プロセッサと、から構成したことを特徴とするネットワーク・インタフェースにある。

【0029】この発明の第20の発明は、データとチャンネルを示す見出し情報とを含むパケットをネットワーク・インタフェースによって処理する方法であって、第1メモリに各通信チャンネル毎の複数の空きバッファ表示を記憶するステップと、パケットを受信するステップと、見出し情報からパケットの通信チャンネルを決定するステップと、パケットの予測サイズに応じて決定されたチャンネル用にメモリから空きバッファ情報の一部分を検索し、その部分を第2メモリに記憶するステップと、第2メモリ内の検索された表示から使用されるバッファを選択するステップと、パケットからのデータを第2メモリ内に記憶するステップと、からなることを特徴とするネットワーク・インタフェースにおけるパケット処理方法にある。

【0030】この発明では上述のように、チャンネル毎の外部の空きバッファ・リストと、入力されるパケットの予測サイズに応じて外部リストの一部分をアクセスすることによって、各々の入力されるパケット用にロードされた対応するチャンネル毎の内部の空きバッファ・リストと、を利用するネットワーク・インタフェースを提供することによって、従来技術に伴う問題点を克服する。

【0031】“チャンネル毎の空きバッファ・リスト”と

いう用語は特定の設立された通信チャンネルにバッファが割当てられることを意味し、それによって、他のチャンネルをバッファに書き込むことができないので、アプリケーションスペースで各バッファを有効に利用できる。この構成によって通信の待ち時間が短縮される。さらに、チャンネル毎の空きバッファ・リスト内のバッファを適宜に制御することによって、アプリケーションスペースの指定位置にデータを配することができる。その上、付加的に唯一必要なことは、チャンネル毎の空きバッファ・キュー(待ち行列)に空きバッファを追加するためのオペレーティング・システムの呼び出しだけである。ネットワーク・インタフェースは広域的空きバッファ・キュー(待ち行列)とチャンネル毎の空きバッファ・キュー(待ち行列)との双方を支援できるので、アプリケーションはその通信要求に基づいて方式を選択することができる。

【0032】一実施の形態におけるネットワーク・インタフェースは、入力されるパケットを処理し、パケット内のセルからのデータを指定チャンネル用の空きバッファへと転送するパケット・プロセッサを含んでいる。パケット・プロセッサは各通信チャソル用の内部の空きバッファ・リストを記憶するメモリを有している。内部の空きバッファ・リストは、各通信チャンネル用のより大量の空きバッファ・リストを記憶している外部の空きバッファ・リスト・メモリからロードされる。所定のチャンネル用のパケットが到着する毎に、パケット・プロセッサはそのチャンネル用の外部の空きバッファ・リストの一部分を検索し、その部分を上記チャンネル用の内部の空きバッファ・リストへとロードする。

【0033】パケット・データが到着すると、パケット・データは内部リスト内のバッファに蓄積される。パケットを処理した後にバッファが残っている場合は、それらは廃棄される。内部の空きバッファ・リストを使用し尽くした後でパケット・データが残っている場合は、追加の空きバッファを要求する割り込みが行われる。割り込みはオペレーティング・システム、または別のアプリケーションプログラムへと向けられる。外部の空きバッファ・リストは空きバッファ・リストの結合リストのような二次元構造として編成できる。チャンネル用の外部の空きバッファ・リストのアプリケーションプログラムをチャンネル毎に制御することによって、アプリケーションプログラムはどこにデータを蓄積するかを制御することができる。

【0034】例えば、アプリケーションプログラムはメモリの任意のサイズの連続部分を外部バッファ・リスト内の空きバッファとして指定することができる。外部バッファ・リストが空になる前に割り込みが強制的に行われるように、割り込みコードを外部バッファ・リストにも与えることができる。この発明を利用することによって、ネットワーク・インタフェースはホスト・オペレーティング・システムとの対話をしなくても、パケット・



11

データをアプリケーションレベルのメモリに容易に蓄積することができる。

【0035】この発明はあらゆる種類のパケットを利用した通信用のネットワーク・インタフェースに該当するものである。パケットを利用した通信の特定の種類の一つは非同期転送モード(ATM)通信、すなわち接続を利用した通信規格であり、これは各通信チャネル毎に明示接続が確立されるということを意味している。ATM通信システムでは、基本ユニットは固定サイズのセルであり、サイズは53バイト、データ・ペイロードは48バイトである。パケットは送信側でセルへと細分化され、ネットワークを通してセルとして伝送され、次に受信側でパケットへと再構成される。

【0036】

【発明の実施の形態】以下、この発明の実施の形態を図面を参照して以下に詳細に説明する。

【0037】ここで図5を参照すると、代表的な非同期転送モード(ATM)通信システムはネットワーク・インタフェース30を有し、パケット32はこれを介して受信されている。パケット32は一般に複数のセル34からなっている。一般に、ネットワーク・インタフェース30はホスト・プロセッサ(図示せず)に接続され、このホスト・プロセッサは、バッファ41を含む関連するホスト・メモリ40を有するオペレーティング・システム38の制御のもとで動作する。このホスト・プロセッサを使用するアプリケーション46および48のような別のアプリケーションは、関連する独自のアプリケーションレベル・メモリ42を有している。ネットワーク・インタフェース30は制御線44を介してオペレーティング・システム38と通信し、データ線45を経てデータを送送する。

【0038】前述したように、一般に図5に示したようなシステムでは、ネットワーク・インタフェース30はパケット32を受信すると制御線44を介してホスト・プロセッサおよびオペレーティング・システム38に割込む。ネットワーク・インタフェース30は各セルからのデータをホスト・メモリ40内の空きバッファへと複写する。その後、データはセル34内の見出し情報によって表示されるアプリケーションのアプリケーションレベル・メモリ42内の記憶位置へと複写され、またはデータの記憶位置がマップされる。

【0039】この従来のシステムは、図6により詳細に図示してある。すなわち、オペレーティング・システム38は空きバッファ・リスト43と、使用中バッファ・リスト49と、通信毎の使用バッファ・テーブル47とを保持している。空きバッファ・リスト43はバッファ・プールすなわちホスト・メモリ40内の利用できる空きバッファの待ち行列(キュー)である。パケットが受信されると、線44aを介して次の空きバッファへのポインタが検索(読み出し)され、データがバッファに書き

12

込まれる。そのバッファへのポインタは線44bを介して使用中バッファ・リスト49へと書き込まれる。通信毎の使用バッファ・テーブル47は、特定のアプリケーションに利用される全ての使用中バッファを表示する。アプリケーションが読み取りを行う場合、そのアプリケーションは通信毎のバッファ・テーブル47中の次の使用中バッファへのポインタを検索し、バッファ内のデータをアプリケーションレベル・メモリ42内に複写、もしくはマップする。

【0040】これに対してこの発明では、空きバッファを表示するリストが各チャネル毎に備えられている。図1は、この発明の一実施の形態によるネットワーク・インタフェースに従ったシステムの構成を示す図である。パケットが受信されると、リストからパケットが選択され、データはオペレーティング・システム・レベルのバッファに複写されずにそのバッファ内に蓄積される。すなわち、各々のアプリケーション46、48はそれぞれ独自の空きバッファ・リスト60、62を有しており、これらはアプリケーションレベル・メモリ42、オペレーティング・システム38、またはネットワーク・インタフェース31内に記憶しておくことができる。

【0041】ネットワーク・インタフェース31は各パケット毎に、受信されたパケットからのデータが蓄積される次のバッファを特定する。このようにして、データはオペレーティング・システム38のバッファには複写されない。さらに、アプリケーションはデータが蓄積される空きバッファを制御することができ、アプリケーションがバッファ内のデータ編成をより厳密に制御することが可能になる。例えば、バッファはメモリ・ページのサイズに対応する必要がなく、所望のとおり小さいサイズでも、大きいサイズでもよい。加えて、アプリケーションは一連の直列バッファを特定することができる。

【0042】図2は図1のネットワーク・インタフェースをより詳細に示している。このネットワーク・インタフェース31はパケット・プロセッサ50とバッファ・リスト52とを含んでいる。パケット・プロセッサ50は各々の通信毎に、この通信に割当てられた全ての空きバッファ記憶位置をリストした空きバッファ・テーブル54の表を含んでいる。パケット・プロセッサ50が単一の集積回路である場合、このテーブルは内部のテーブルである。ある通信用に入力されるメッセージは、オペレーティング・システム・メモリへと送られるのではなく、アプリケーションレベル・メモリ42内(ホスト・メモリで構成される場合もある)のその通信用の空きバッファ・リスト内に見い出された記憶位置に直接複写される。空きバッファが発見されない場合は、割り込みが行われ、アプリケーションまたはオペレーティング・システムが利用できる空きバッファの新たなリストを作成し、それらをバッファ・リスト52へと追加する。

【0043】パケット・プロセッサ50がパケット32

を受信すると、ネットワーク・インタフェース31は、空きバッファ表示が内部でロードされることを保証する。ネットワーク・インタフェースは割り込みコードを作成することもできる。空きバッファ表示が未だロードされていない場合は、空きバッファ表示は外部リストからロードされる。外部リストが空である場合は、ネットワーク・インタフェースはサービス用の割り込みコードを発生する。パケットの処理中は、パケットの終端が受信されるまで、またはロードされたリスト内のバッファが使い果たされるまで、バッファにはパケット・データが充填される。後者の場合は、追加のバッファがネットワーク・インタフェースにロードされ、そのどれもが使用できない場合は割り込みが行われる。ネットワーク・インタフェースがパケットの処理を終了すると、おそらくは割り込みを利用してアプリケーションにその旨の信号が送られることが一般的である。

【0044】バッファ・リストを作成し、パケット・プロセッサ50によってバッファを利用するには多くの異なる方法がある。

【0045】バッファ・リストを利用する第1の方法には、パケットを受信した時点で、バッファ・リスト全体を内部の空きバッファ・テーブル54内に複写することが含まれている。ある通信についてパケットの処理が終了すると、バッファ・リストの未使用部分は廃棄される。この解決方法の問題点の一つは、大量の空きバッファ・リストが廃棄されることがあり、新たなリストを作成しなければならず、ローカルプロセッサ、またはホスト・プロセッサによる処理時間が必要になることである。

【0046】第2の方法もバッファ・リスト全体を内部の空きバッファ・テーブル54内に複写する。しかし、パケットの処理が完了すると、バッファ・リストの未使用部分は廃棄されず、内部リスト内に保存される。この方法の問題点の一つは、自己洗浄作用がないことである。すなわち、ローカルまたはホスト・プロセッサは以前のパケットから残されたバッファの取り出し処理を行わなければならない。このような取り出し処理は、以前のパケットで使用されたバッファとは異なる特定のバッファ内に次のパケットを記憶しておくべき場合に必要となる。入力されるパケットのサイズは一般に事前には判明していないので、この第2の方法ではパケットの浄化は常に必要となる。このような浄化処理はクリティカルパス(限界経路方式)であり、性能を損ねるばかりではなく、パケット・プロセッサをこのような浄化を取り扱うように設計しなければならない。

【0047】上述の解決方法ではいずれも、一連のパケットを連続的な集合としてアプリケーションレベルのメモリに送る能力を与えるのが困難である。すなわち、第1のパケットの後で、第2のパケットを連続するバッファ記憶位置に記憶しなければならない場合、上述の2つ

の解決方法ではそのような能力を与えることができない。

【0048】入力されるパケットのサイズは事前には判明していないものの、これは限定できる場合が多く、空きバッファのリスト内のバッファ容量は上述の限定に合わせて設定することができる。従って、主要な問題は上記の限定と、実際のパケット・サイズとの差に起因する残されたバッファに関連する問題である。場合によっては、限定を判断することが困難な場合もある。この場合、ある範囲の限界を設定することができ、新たな空きバッファ・リストをロードしようとしてその限界を超えた場合、次の空きバッファ・リストを使用でき、または、そのリストが不適切であるか、または存在しない場合は、ローカルまたはホスト・プロセッサが追加の空きバッファを指定するように割り込みを発生することができる。このような洞察が第3の方法の基礎になっている。

【0049】次に図3を参照して、バッファ・リストをより厳密に制御できる第3の方法を説明する。この方法では、外部バッファ・リストの一部分だけが内部空きバッファ・テーブルへと読み込まれる。読み込まれるバッファの数は入力されるパケットの予測サイズに応じて選択される。このような能力を付与するために、この発明の一実施の形態では、バッファ・リストは二次元構造にされており、複数の結合されたリストの結合リストであるとみなすことができる。すなわち、空きバッファ・リストが作成され、複数のリストへと分解されるのである。各リストの長さは読み込まれるバッファの数に対応する。例えば、空きバッファの各々の表示は1つの予約された領域かまたはいずれかの次の空きバッファへのポインタとして使用される別の領域および、いずれかの次の空きバッファ・リストへとリンクするためのさらに別の領域を含み得る。

【0050】パケット・プロセッサ50はパケット32を受信すると各通信毎のいずれかの次のバッファ・リストの記憶位置を記憶する。記憶されたポインタが存在しない場合は、空きバッファのリストが作成されるように割り込みが行われる。その後、バッファ・リストは一時レジスタ内に記憶された次のバッファ・リストへのポインタとともに検索される。あるいは、次のポインタは第1のポインタからの定常偏差でもよく、これは単に計算すればよい。次のバッファ・リストは最初のバッファ・リストが無くなったときだけ検索される。最初のバッファ・リストはメモリ内に保持しておくことができ、後続のいずれかのパケットを受信した時点で廃棄し、またはパケットの処理が完了したときに廃棄すればよい。

【0051】この方式の利点の一つは、ローカルまたはホスト・プロセッサがバッファ・リストの管理の複雑さをほとんど処理できることにある。すなわち、負担の殆どはソフトウェアにある。しかし、二次元構造のバッ

10

20

30

40

50

15

ァ・リストを使用することにより、上述の第1と第2の方法のいずれをも実行するフレキシビリティが得られる。外部空きバッファに新たな空きバッファ・リストを追加することには潜在的な矛盾が生ずる。すなわち、バケット・プロセッサはローカル・プロセッサがリストを修正すると同時にリストを修正することがあり得る。

【0052】この矛盾を解決するため、チャンネル毎の保持および保持解消(hold and unhold)動作を利用できる。リストを修正する前に、プロセッサはそのチャンネルを“保持”し、次に修正が終了した後、チャンネルの“保持を解消”する。チャンネルが“保持”状態にあるときに到着するデータの損失を防止するため、適当な緩衝処理を行わなければならない。あるいは、プロセッサが先ず、利用できる空きバッファ・リストの数をチェックし、複数の空きバッファ・リストのリストが一つ、または2つの空きバッファ・リストしか含んでいない場合には“保持”および“保持解消”動作を行うことができる。上記の数以上の空きバッファ・リストを利用できる場合は、プロセッサは、外部リストがバケット・プロセッサによって使用し尽くされる前に、新たな空きバッファ・リストを外部リストの端に追加することができる。

【0053】次に新たな空きバッファ・リストを作成するための割り込みの制御を図4を参照して説明する。一般に、空きバッファ・リストをそれ以上利用できない場合に割り込み動作を行うことができる。加えて、最後の空きバッファ・リストでは、複数の空きバッファ・リストのリスト内の次の空きバッファ・リストへのポインタは常に存在しない。従って、バケット・プロセッサは次の空きバッファ・リスト・ポインタをチェックし、そのポインタが存在しない場合だけ割り込みコードを作成する。

【0054】さらに補足的な方法は、空きバッファ・リストのリンクされたリスト表現の各々のバッファ・リストにコードを追加することである。例えば、複数のバッファ・リストのリンクされたリスト内の要素を表す図4に示すデータ語70は、参照番号72で示すように、次の空きバッファの記憶位置を記憶する記憶位置へのポインタを含んでいる。さらに、74で示した次のバッファ・リストを記憶するアドレス位置にポインタが与えられる。そこで割り込みコード76が現在のバッファ・リストのために付与される。この割り込みコードは、別の空きバッファ・リストを作成するための割り込みを発生するか否かを判定するバケット・プロセッサによって評価される。

【0055】このようにして、ネットワーク・インタフェースが入力されるバケットを処理している間に、ローカルまたはホスト・プロセッサはより多くの空きバッファ・リストを発生することができる。各空きバッファ・リストの見出しに割り込みコードを適宜に設定することによって、特定の複数の空きバッファ・リストのロード

16

中に、または最後の空きバッファ・リストのロード中に、全ての新たな空きバッファ・リストをロードするために割り込みを発生することができる。この割り込みコードによって、アプリケーションはそれまでにいくつの空きバッファおよび空きバッファ・リストが使用され、空きバッファと空きバッファ・リストがいくつ残っているかを判定することができる。この割り込みはさらに、特定の空きバッファの集合がロードされた時に処理を実行することができる。最後に、データ語70は未使用領域78を含んでいてもよい。

【0056】以上のようにチャンネル毎の空きバッファ・リストを使用したこの発明によるネットワーク・インタフェースは、入力されるバケットを処理し、バケット・データをバケットが属するチャンネル用に指定されたバッファ内に記憶するバケット・プロセッサを含んでいる。このバケット・プロセッサは各々の通信チャンネル用に内部の空きバッファ・リストを記憶するメモリを有している。内部空きバッファ・リストは、各通信チャンネル用により大量の空きバッファ・リストを含んでいる外部の空きバッファ・リスト・メモリからロードされる。所定のチャンネル用のバケットが到着する毎に、バケット・プロセッサはそのチャンネル用の外部空きバッファ・リストの一部分を検索し、その部分を当該チャンネル用の内部空きバッファ・リストへとロードする。ロードされた部分は、入力されるバケットの予測サイズを処理するのに充分であるとみなされる多数の空きバッファである。

【0057】バケットの処理がなされると、データは内部リストのバッファ内に蓄積される。バケットが処理された後で未使用の内部バッファはいずれも戻された使用中バッファのリストの終端に付加され、自己洗浄特性を有し、それによってアプリケーションは、入力されるバケットがどのバッファを利用するかについてのある種の制御を行うことが可能になる。内部の空きバッファ・リストが使用し尽くされた後にバケット・データが残っている場合は、バケット・プロセッサは外部空きバッファ・リストの別の部分を検索し、バケットの処理動作を継続する。外部空きバッファ・リストにそれ以上バッファが残っていない場合は、追加の空きバッファを要求するように割り込み(コード)が発生される。この割り込みはネットワーク・インタフェース、またはホスト・プロセッサと連結されているローカルプロセッサに向けられる。外部空きバッファ・リストの一部だけを読み取ることを促進するため、リストは複数の空きバッファ・リストの結合リストのような二次元構造として編成することができる。

【0058】チャンネル用の外部の空きバッファ・リストのアプリケーションプログラムをチャンネル毎に制御することによって、アプリケーションプログラムはどこにデータを蓄積するかを制御することができる。例えば、アプリケーションプログラムはメモリの任意のサイズの連

17

続部分を外部バッファ・リスト内の空きバッファとして指定することができる。外部バッファ・リストが空になる前に割り込みが強制的に行われるように、割り込みコードを外部バッファ・リストにも与えることができる。

【0059】これまでこの発明の実施の形態を説明してきたが、これまでの記述は単に説明目的のためであり、限定されるものではなく、例示したものであるに過ぎないことが専門家には明白であろう。専門家には多くの修正と別の実施の形態の実現が可能である。例えば、外部バッファ・リストは一つのリストでもよく、また、パケット・プロセッサは外部バッファ・リストから読みだされた多数のバッファを表示するプログラム可能なレジスタを含んでいてもよい。これらの、およびその他の修正は添付の特許請求の範囲に含まれるものである。

#### 【図面の簡単な説明】

【図1】 この発明の一実施の形態によるネットワーク・インタフェースに従ったシステムの構成を示す図である。

【図2】 図2のネットワーク・インタフェースをより詳細に説明するための図である。

18

【図3】 チャンネル毎の外部空きバッファ・リストの好ましい構造を示す図である。

【図4】 割り込みコードを含む連結されたリスト・データ構造の一例を示す図である。

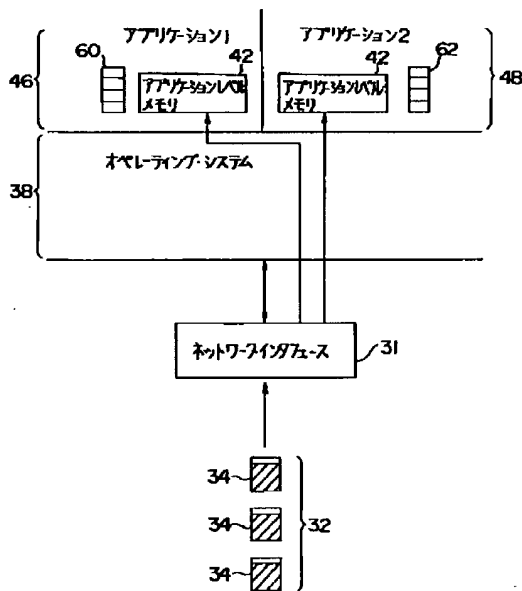
【図5】 広域的空きバッファ・リストを使用した多重通信(結線)用の入力されるパケットを処理するための従来のネットワーク・インタフェース・システムを示す図である。

【図6】 図5のより詳細な図である。

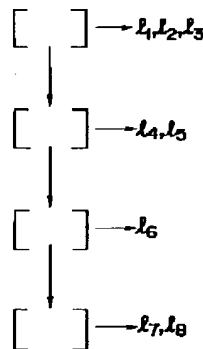
#### 【符号の説明】

30, 31 ネットワーク・インタフェース、32 パケット、34 セル、36 ホスト・プロセッサ、38 オペレーティング・システム、40 ホスト・メモリ、41 バッファ、42 アプリケーションレベル・メモリ、43 空きバッファ・リスト、44 制御線、45 データ線、46, 48 アプリケーション、47, 49 使用中バッファ・テーブル、50 パケット・プロセッサ、52 バッファ・リスト、54 空きバッファ・テーブル、60, 62 空きバッファ・リスト。

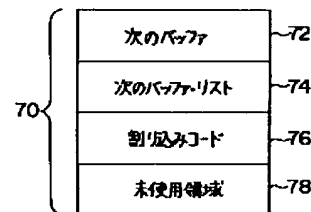
【図1】



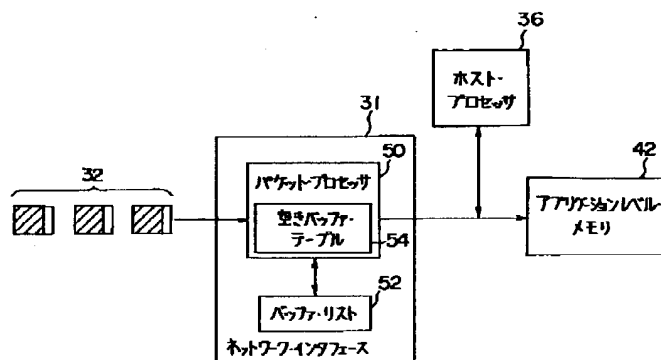
【図3】



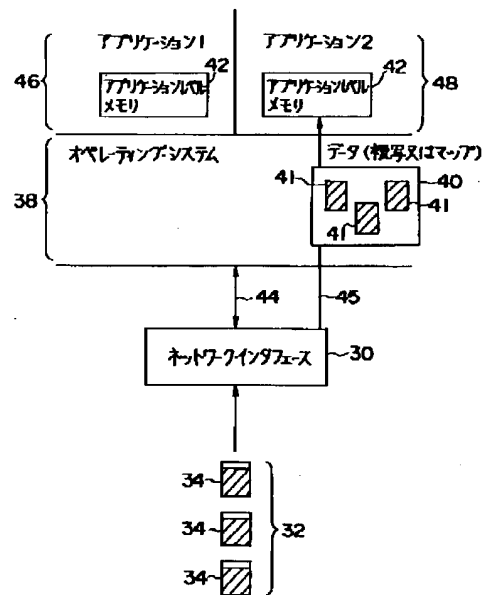
【図4】



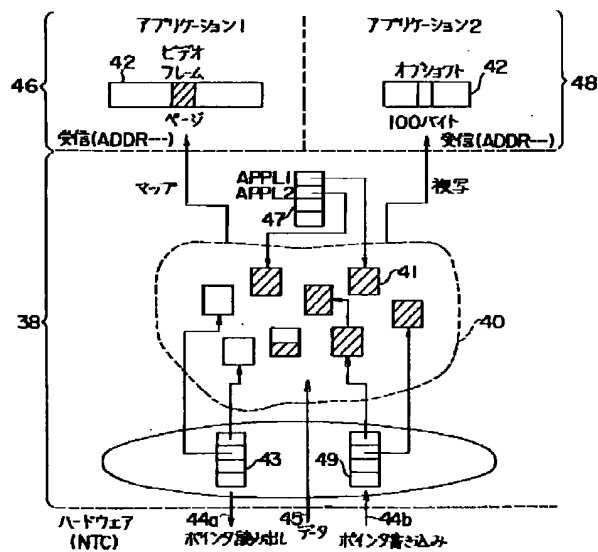
【図2】



【図5】



【図6】



フロントページの続き

(71)出願人 595151497  
201 BROADWAY, CAMBRIDGE, MASSACHUSETTS  
02139, U. S. A.